Lexical Selection in Language Production

Winston Downes

A thesis submitted to the Victoria University of Wellington in fulfillment of the requirements of the degree of Master of Science in Psychology.

Victoria University of Wellington 2011

Two experiments were conducted with first-year university students in an effort to discover more about what happens when a phrase is spoken. A paradigm was constructed with the intention of getting the participants to produce a simple, two-noun phrase at a cue and then ‘catch’ them out having them say the name of a single picture presented instead. The single picture presented to ‘catch’ the participants out (instead of the cue) was either the first or second name in the simple two-noun phrase, or a third, unplanned picture. The intention was to compare the relative timings of the different catch pictures in an effort to discover which of two theories of speech production best describes the cognitive processes that underlie such processes. The second experiment was an extension of this idea but also included a semantic relatedness variable, where the catch picture could be semantically related to an item shown during the planning of the simple, two-noun phrase. The results of these experiments were not in line with the hypothesis regarding the relative timings of the catch pictures, but were in line with the hypothesis that it would take longer to name catch pictures that were preceded by semantically related pictures. Implications of such findings are discussed along with possible future modifications to extend the utility of the paradigm used in this study.

Acknowledgements

I would like to thank my supervisor Carolyn Wilshire, for her methodological guidance and extraordinary patience throughout the completion of my thesis; my parents for their ongoing support and my dear friend Rikka, who encouraged me throughout my study.

Table of Contents

Tables and figures

Figures

Tables

Lexical Selection in Language Production

Stringing together a sentence is an ability that most of us take for granted, and the ease at which this typically occurs belies the cognitive complexity that this process entails. When we produce a sentence, we must not only retrieve information about each of the words within it, but we must also time their production so that they are uttered in the correct order. Not only would understanding how we are able to produce sentences be illuminating as far as cognitive theory is concerned, but it may help us understand what happens when deficits in speech production occur. For example, in disorders such as non-fluent aphasia, which is characterised by effortful, halting speech, sufferers have particular difficulty in producing phrases and sentences, especially when the phrases are complex and/or the nouns are semantically related (Freedman et al. 2004; Wilshire & McCarthy, 2002; Scott & Wilshire, 2008).

This overview of the literature first considers contemporary theories of single word production. It then discusses research that focuses on longer utterances and mechanisms that may play a part in forming these utterances. It then sets out the rationale underlying the current study.

Theories of Single Word Production

A widely held assumption in the cognitive psychology literature is that word production involves two stages (Dell, 1986; MacKay, 1987; Badecker et al., 1995; Rapp & Goldrick, 2000; Schwartz, Dell, Martin, Gahl & Sobel, 2006). For example, when one has to name a picture, the first stage involves the selection of the appropriate lexical unit from its semantic representation (sometimes referred to as conceptual semantics or message level) of the picture to be named. There is dispute over what a lexical unit typically entails. Badecker et al. state that it is a record that contains the semantic and grammatical features of an item, but not those of its orthography and phonology (See Freedman et al., 2004 for an alternative definition). The second stage involves the selection of the phonological form of the word. Many network models incorporate this two-stage model but represent the semantic, lexical and phonological information in particular nodes that are linked to each other. There are typically three primary levels of nodes: semantic, lexical and phonological (see Mackay, 1987; Dell, 1986; Freedman et al., 2004 for examples of network models).

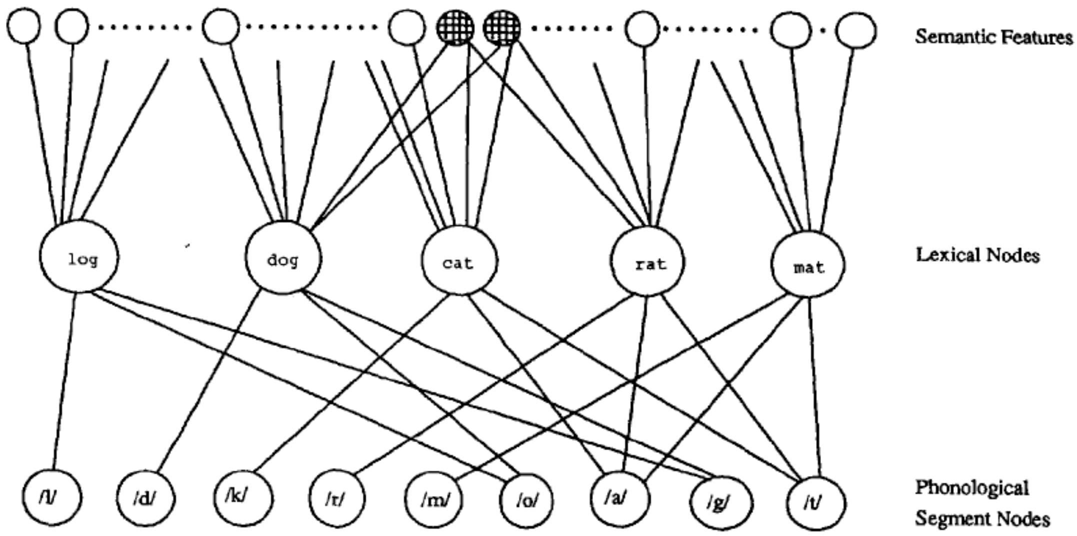

Figure 1. Dell's lexical network spreading activation model for speech production.

An example framework for illustration is Dell's spreading activation model (Dell and O'Seaghdha, 1991; Figure 1). In this model there are three levels of representation: semantic, lexical and phonological. The semantic level consists of units (or nodes) that represent semantic features of an item that is being processed; the lexical, or lemma level, consists of units corresponding to each word represented in our stored knowledge; and the phonological level consists of units representing each of the phonemes in the language. When a picture is to be named, the features that comprise it activate the corresponding units within the semantic level, which then activate the appropriate lexical units through their connections. When a single lexical unit reaches a certain level of activation (or an activation level above other units by some threshold), it receives an additional boost to its activation levels. Activation flows from the lexical units to the corresponding phonological units, which leads to the picture being named. Activation within the network is bi-directional so, for example, activation of the lexical units would send activation to both the semantic and phonological units and phonological units can feed activation back to lexical units. Also, this activation is automatic or unconstrained and when activation of a unit ceases, then the unit does not stay activated at the same level until further input but attenuates; that is, there is decay in activation.

A crucial step in this process is how we ensure that any one lexical unit is selected for production. Consider, for example, the task of producing a single word in isolation when naming a picture presented alone, such as a picture of a cat. Within the lexicon and within the semantic level, units that comprise the cat's features would become activated (e.g. fur, whiskers); however, there are a few other items that share certain features with the cat, for example, dogs or other feline species such as tigers and panthers. With a spreading activation model such as Dell and O'Seaghdha's (1991), the lexical representations of these items would also receive some activation. However, since “cat” shares more features than any other item, it will receive the most activation. Furthermore, it will also receive an additional boost to its activation levels when it is “selected” for production.

Other theories postulate slightly different mechanisms for ensuring that the most highly activated lexical item has a sufficient leading edge. For example, Wheeldon and Monsell (1994) have suggested lateral inhibition, a mechanism by which lexical units inhibit other units in proportion to their own levels of activation. Consequently, the more activated an item is, the more it will inhibit its ‘competitors’. So in our cat example above, the lexical unit for “cat” inhibits the other competing lexical units as it becomes activated, such as “dog” and “tiger”. All lexical units are capable of this inhibition and the most activated will inhibit the most strongly (see also McClelland & Elman, 1986; McClelland and Rumelhart, 1981).

Another type of mechanism is proposed by Wilshire and McCarthy (2002), who posit a mechanism external to the lexicon, which operates by modulating activation levels of lexical items according to the current task requirements. Much of the supporting evidence for this proposal comes from aphasic individuals. In a study with a non-fluent aphasic patient (BM), Wilshire and McCarthy found that BM’s picture naming accuracy was profoundly affected by contextual manipulations, such as increasing the presentation rate and/or the semantic similarity of the pictures to be named. In a cyclic naming task, where a small series of items were presented repeatedly, each time in a different random order, BM’s accuracy dropped when the pictures in the set were semantically related (e.g. orange, apple, banana, grapes, lemon, pear), and the magnitude of this “semantic blocking” effect increased with the rate of presentation. Many of BM's errors were perseverations of other previously named items from the same set. Using Dell's aforementioned lexical model as a framework, (see Figure 1), Wilshire and McCarthy posited that BM had sustained damage to a mechanism outside of his lexicon that modulates and controls activation within it, enabling the desired word to be effectively selected form amongst other activated competitor words. By this view, BM’s particular difficulty with semantically related word arises because these words activate one another via their shared features.

Therefore, semantically related words compete more strongly with the target than unrelated ones. Here, the lexicon itself is working as it should: activation spreads as would be expected in intact normals; however, an extra-lexicon control mechanism is posited to be compromised.

A similar idea was advanced by Thompson-Schill and Botvinick (2006), who also advocate a mechanism that responds to the current task demands. These authors hypothesise that when one is presented with a stimulus to which a response must be made, a pattern of activation occurs over several possible relevant responses. Another process then translates this probability distribution into a single response. However, the task itself constrains the responses that will be made. The authors argue that a top-down mechanism (that is, a mechanism not driven by the stimulus itself) changes the activation weights among the possible responses which biases responses to be made that are appropriate to the task at hand. For example, in a verb generation task, if the word 'canoe' is given as a stimulus, the participant will be required to name verbs that come to mind. Among the responses that may begin to show automatic activation from such a word (e.g. swim, boat, row, float, water), Thompson-Schill and Botvinick posit that there exists influence from a top-down mechanism that biases responses towards the verbs (e.g. swim, row, float).

Production of Words in Context

Of course, words are not usually produced in isolation, but rather in the context of longer utterances. Under these circumstances, there may be activation not only involving the target and its closest semantic neighbours, but also involving other items planned for the same utterance. Here, an additional mechanism may be required to ensure the correct item is produced at the right time. To explore the specific mechanism that might be engaged in this situation, researchers have often used a multiple picture naming tasks, in which a series of pictures must be named in a particular order, often as part of a phrase or sentence. For example, in an experiment in which participants were required to name (in order) two objects that appeared on screen, Freedman et al. found that controls took longer to initiate the phrase when the objects were semantically related (e.g. nose and mouth) than when they were unrelated (nose and hat). Again, evidence from individuals with nonfluent aphasia suggests that production of words in context may engage specific processes that can be selectively impaired after brain damage. Some such individuals with nonfluent aphasia have been found to show greatly exaggerated semantic relatedness effects on asks involving the production of pairs of picture names (Freedman et al., 1994; Scott & Wilshire, 2008). This evidence has led some researchers to conclude that there exists a mechanism that ensures the correct selection and sequencing of items in multi-word utterances, which has become compromised in aphasic patients.

There are many theories as to how this sequencing mechanism operates. Dell himself argues that there is a syntactic frame containing slots which may be filled with the most highly activated lexical unit of the appropriate grammatical type (e.g., a "noun" slot is filled by the most highly activated noun) and the unit "granted current node status (has) its activation level … boosted" (pp. 289; 1986). So, for example, in the sentence "the boy swims" the lexical unit for "boy" receives a boost when the noun slot in the syntactic frame becomes currently activated, then the lexical unit for "swim" receives a boost when the verb slot within the syntactic frame becomes activated.

Another idea, proposed by Freedman et al. (2004) is that there is a dedicated short-term memory buffer that maintains and supports selection of lexical items in multi-word utterances. They argue that lateral inhibition is not an acceptable explanation for how competition is resolved in this situation because it would have the effect of inhibiting the second item beneath baseline activation making it unavailable for subsequent activation. This account therefore has difficulty explaining the evidence that the second word in a two-noun phrase can influence latencies to produce the first word, if it is semantically related to it (e.g. nose and mouth). In this situation, two words activate one another via their semantic features, and this makes it harder for the first lexical unit to become activated above that of the second by the critical threshold. As a consequence, the time taken for the activation of the first lexical unit to accrue to a critical level is longer than it would have been had the two nouns been unrelated.

These authors describe a model already familiar, consisting of a network with three levels of units: semantic, lexical and phonological. However, the model also includes a "lexical-semantic buffer", which is a part of short-term memory (memory that holds information for short periods of time while operations are undertaken) and is used to maintain lexical representations. Lexical units within the network have links to this buffer. During word production, semantic units activate their corresponding lexical units which in turn activate particular units within the lexical-semantic STM buffer. When a simple two-noun phrase is to be produced, attention is directed towards the first unit within the buffer (representing the first noun) which, when it becomes sufficiently activated above the other units in the buffer, then feeds activation back through to its original lexical unit. When the two nouns are semantically related then lexical competitors will activate nodes within the buffer as well, which will extend the time taken for the target word to become activated to the critical proportion above that of the competitor.

Freedman et al. are advocating a mechanism that is external to the lexicon (i.e.

lexical-semantic STM) which works a bit like a syntactic frame. This mechanism does not work by inhibiting competitors but by acting as a 'space' where a phrase is planned and, interestingly, it is attention that increases the activation of the target word within the buffer and ensures it activates the corresponding phonological units. The authors suggest that some aphasic patients suffer from rapid decay in the linkages between the lexical units and their corresponding units within the buffer. As a consequence, the STM buffer takes longer to boost the activation of their corresponding lexical items and as a result of the delay there is greater opportunity for activation to spread amongst semantically related lexical units.

A somewhat different proposal, which bears some similarities to that of Wilshire and McCarthy (2002) and Thompson-Schill and Botvinick (2006) was advanced by Biegler et al. (2008). These researchers also advocate a mechanism outside of the lexicon that performs a selective process. However, its primary purpose, they argue, is to inhibit lexical representations once they have been activated. This inhibitory mechanism, they argue, operates specifically at the lexical level. They report findings from a non-fluent aphasic patient (M.L.) who, similar to Wilshire and McCarthy's patient BM, showed a pathologically marked semantic blocking effect when asked to repeatedly name a small series of pictures. However, when the task did not require naming a picture, but only matching it to a name provided, M.L. showed no such abnormalities: he performed like controls. Further, M.L. also performed poorly on the Recent-Negatives task, in which he had to view a small list of three words (one after the other), and then indicate whether a fourth (the probe) appeared in the list or not (Hamilton & Martin, 2005). Negative trials were those where the probe did not appear on the list but did appear in a previous one. On this task, M.L. showed a tendency to indicate that the probe had indeed appeared in the present trial which suggests susceptibility to interference from previous trials. This suggests that the trouble experienced by non-fluent Aphasics where multi-word utterances are concerned is that they suffer interference from previously accessed lexical units when they must continue to access more units to continue to speak. As indicated above, semantic contexts would tax the lexicon more; however, M.L.'s performance in Hamilton and Martin's (2005) Recent-Negatives task suggests that interference can occur even in non-semantic contexts.

This result seems paradoxical at first, because a STM patient with a memory span of 2.5 items would be expected to show little interference from previous trials. The fact that M.L. did show interference from previously shown items, even ones from three lists back, suggests that M.L. may suffer from a difficulty in inhibiting persisting activation from previously activated items, not from a deficiency in retaining them. (See Barde et al. (2010) for an alternative account of such findings based not on inhibition, but rather on weak STM, as defined as reactivated LTM).

These inhibition accounts tend to consider inhibition as a single unitary mechanism, regardless of whether it involves inhibition of previously processed material, or material planned for future use. However, some other researchers have distinguished between different types of inhibition. For example, May et al. (1999) argue that as far as online processing is concerned, inhibition serves three purposes, which are: (1) the restriction of activation to relevant items only, (2) the deletion of items that are no longer relevant and (3) the restraining of highly probable responses to enable a less probable response to be activated. According to this framework, proactive interference tasks, such as the Recent-Negatives task and semantic blocking paradigms would tap the deletion process (2), whereas interference paradigms such as the Stroop task would tap the restraint process (3).

Nevertheless, some recent findings from individuals with nonfluent aphasia appear to support the idea of a single common mechanism. Hamilton and Martin's patient M.L was found to be impaired not only on tasks requiring him to delete previously relevant responses, such the semantically blocked naming task, and the recent- negatives task, but also those that involve restraining an immediately available response such as the Stroop colour-word task. More recently, Scott and Wilshire (2008), reported a Broca's aphasia patient (J.H.M.) who exhibited specific abnormalities on three quite different tasks: First, on a picture pair production task, JHM was significantly and abnormally slower to initiate her response when the target items were semantically related (e.g., goat and pig) than when they were unrelated (goat and lamp). Second, on a semantic blocking task, JHM showed exaggerated effects of semantic relatedness amongst the target pictures. And third, on the Stroop colour-word task, JHM naming times were abnormally slowed when the name of the word being viewed conflicted with its ink colour. It is hypothesised that there was an impairment to a selection and control mechanism which would have normally acted to dampen the activation of the relevant competing items, whether they be ones that have been previously produced (such as in the semantic blocking task), ones that are planned for upcoming production (such as in the picture pair naming task), or ones that are preferentially activated by the stimulus items (such as in the Stroop task). Therefore, it may be that the type of inhibition required in all these tasks is the same, and that the difference lies in paradigm only. This explanation is in line with Hamilton and Martin's (2005) and Biegler et al.'s (2008).

The Current Study

The focus of the current study is on the processes that are involved in sequencing words within short multi-word utterances. The theories in the above review explain the word sequencing process in two major ways: a) what we may describe as differential activation created from boosting of the target item over that of other items already produced and/or those planned for later in the utterance (Dell, 1986; Freedman et al., 2004); and b) what we might describe as inhibition or dampening, created when upcoming items are inhibited so as to enable production of the current item (Wilshire and McCarthy, 2002; Hamilton and Martin, 2005; Biegler et al. 2008). In the boosting account of selection, only the target has to be activated, therefore, the target of the mechanism only involves one item. In the inhibition account, the mechanism has to target many possible lexical nodes, such as those that have been previously selected, those that are planned for subsequent production in the same utterance, and those that become activated because they share a semantic relationship to the intended word. Establishing which of these purported mechanisms best describes the process of lexical sequencing will increase our knowledge of the processes engaged in phrase and sentence production. This knowledge may lead to a better understanding of sentence production deficits such as those seen in non-fluent aphasia.

One way to gain further insight into the nature of the sequencing process is to examine the state of the production cycle just prior to production of a phrase (after planning is largely complete, but before initiation of the planned utterance). The basic task employed in the present study is an extension of that used by Freedman et al. (2004) and Scott and Wilshire (2008). A series of two pictures is presented and the participant is subsequently cued to produce them in their original presented order, conjoined with an “and”. However, on 20% of the trials, (“catch” trials), instead of being cued to produce the planned phrase, the participants are instead presented with a single picture which they are required to name. This picture may be either one of the two pictures that were originally presented, or an entirely different, unplanned picture. The catch trials are timed to occur at the moment just prior to initiation of the planned utterance. At this time, it is hypothesised that particular differences in the naming latencies for the catch trial pictures will reveal which of the two mechanisms is used to resolve competition. If a catch trial involving the unplanned picture has a shorter latency than one involving the picture planned for second position, then the inhibition theory will be supported; if the unplanned picture's latency is no different than that of the second, then the boosting theory will be supported.

Experiment 1

In this experiment, participants were required to view two pictures one after the other, and name them in the order that they were seen with “and” in between them. For example, if a picture of a cat was shown on a computer screen, followed by a picture of a tree, then at a prompt the participant would be required to say “cat and tree”. However, this was the procedure only 80% of the time (henceforth known as 'filler trials'). In the other 20% of trials, the procedure was the same up until the point where the prompt would be expected to appear. However, instead of the prompt, the participants saw either the first, second, or a third picture, which was not shown in the trial; the participant was required to name this picture as quickly as possible without compromising accuracy (these trials are henceforth known as 'catch trials'). The filler trials are in the experiment for the sole purpose of ensuring that the target phrase is being actively planned prior to presentation of the response cue. The variable of interest is the latency to produce the pictures during the catch trials.

On these catch trials, naming latencies are predicted to be faster when the picture presented instead of the cue depicts the word planned for first position in the original phrase (e.g., planned phrase: "cat and tree"; catch trial target "cat"; henceforth, the *first-planned* condition), than when it depicts the word planned for second position, (e.g., planned phrase: "cat and tree"; catch trial target "tree"; henceforth, the *second-planned* condition). This is because the lexical unit for this first-planned target, being the first item in the phrase to be produced, should already be activated above its second-planned competitor at the time the catch-trial target is presented. This is a consistent prediction of all theories discussed above.

However, predictions regarding *unplanned* catch trials - when the target is unrelated to either of those previously planned for the phrase (e.g. planned phrase: “cat and tree”; catch trial picture: "shoe") - will depend upon the specific theory adopted. According to "activation boost" theories - where each items planned for the phrase receives an additional boost to its activation just prior to production (e.g., Dell, 1986; Freedman et al., 2004) - naming latencies for unplanned catch trials would not be any faster (and indeed will probably be slower) than those for second-planned catch trials. This is because the first-planned word is receiving the boost and the unplanned word is not engaged in any way during planning. Conversely, according to "inhibition" theories - which propose that items planned for non-current upcoming positions are actively inhibited (Wilshire and McCarthy, 2002; Hamilton and Martin, 2005; Biegler et al. 2008) - then latencies to unplanned catch trials will actually be faster than those to second-planned catch trials. This is because the lexical unit for the word planned for second position may be inhibited below baseline to enable the first-planned word to be produced accurately.

Method

Participants

Participants were 10 first year psychology students studying at Victoria University of mixed sex, though most of them were female (n=7). 17 participants participated in the experiment but there were seven who were not focused enough on the task, made too many mistakes or found the experiment difficult because English was not their first language. These participants’ data was not subject to analysis.

Materials

432 line drawings of typical objects and animals were used, belonging to one of seven semantic categories: people, animals, fruit, vegetables, furniture, utensils and tools. The drawings were from a variety of sources, consisting of: coloured photos, black line drawings, coloured drawings and images created by computer. The pictures were shown on a Macintosh computer and each session was taped using a mini-disc recorder. The pictures were of two types: those that were used as filler trial pictures (288 pictures) and those that were used as in the catch trials (144 pictures; see below in '*Design and procedure*' for details about filler and catch trials). The majority of catch trial pictures had previously been normed by a sample of first year university students and had an agreement rating of at least 80%. All pictures used in catch trials pictures had monosyllabic names. The Kucera-Francis (1967) written frequencies for the catch trial pictures ranged 1-1207 per million with a mean of 80.19 and a standard deviation of 154.07 (Wilson, 1987). There were 36 catch trial *target* pictures but, as per the procedure, these pictures had to be presented with another (remember, the catch picture of the cat was presented with that of a tree in our example above). The 36 catch trial target pictures were presented with 108 other pictures throughout the entire experiment, which themselves appeared with the targets depending on the condition in which the targets were in (a full list of catch trial pictures for Experiment 1 can be found in Table 2 in the Appendix). Also, target pictures and the pictures presented with them did not start with the same letter and were not from the same semantic category. As the filler trials were not analysed in anyway, there was no constraint placed on their word length or number of syllables they contained.

Design and procedure

Each participant was required to complete four testing sessions. The first session was a simple picture naming exercise, which was not part of the actual experiment but was included to familiarise each participant to the pictures used in the subsequent sessions. In this naming session, the participant was shown every picture used in the experiment (432 pictures). Each picture appeared simultaneously with a beep and was shown for 2000 ms, then disappeared from the screen. 500 ms later the next picture appeared. The participant was required to view and name aloud the picture with the first name that came to mind. An experimenter was present during this part of the task. In the event where the participant used an incorrect term, the experimenter would correct them. The naming exercise was split into four blocks, to allow the participant small breaks throughout. Depending on the break time taken, these sessions were normally quite short, lasting some 20-23 minutes.

The subsequent three sessions comprised the actual experiment. Each of these three sessions was composed of 180 trials. 144 of these trials (80%) were filler trials and the remaining 36 were catch trials. In the filler trials, two pictures were presented one after the other and participants were required to name the pictures in the order that they appeared in, with the word "and" in between them (see *Figure 2*). Each trial proceeded as follows: at the beginning of a trial, the first picture appeared for 1000 ms, then there was a 500 ms gap which showed nothing but a white screen. After that, the second picture was shown, which also lasted 1000 ms. 500 ms after the second picture disappeared, a prompt that read "____and____" appeared, and stayed on screen for 2000 ms. The participant was then required to name the pictures in the correct order as quickly as possible without compromising accuracy. In each session, pictures for the filler trials were drawn from the same pool of 288 pictures (two pictures per trial across 144 trials). The order of presentation was different for each session and the pairings for the two pictures were varied across the three experimental sessions. Pairs used in the second experimental session were the same as those used in session 1, except the order of appearance was swapped. For the last session, all filler pictures were re-shuffled and re-paired to create entirely new pairs.

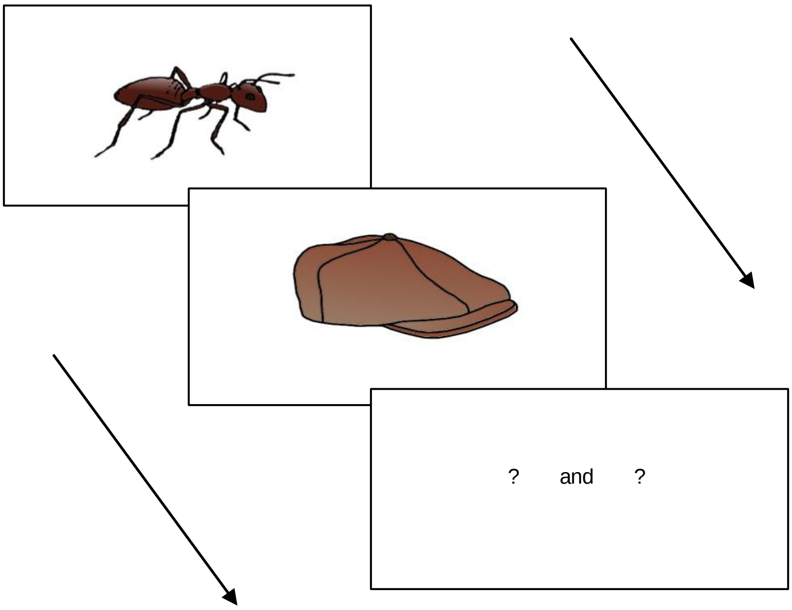

Figure 2. The procedure for the filler trials

The remaining 36 trials in each experimental session (20%) were “catch” trials. The procedure in a catch trial was exactly the same as that mentioned above up until the prompt (____and____). Instead of seeing the prompt, the participants were shown another picture, which was either a repeat of either the first or the second of the pictures just viewed (referred to as first-planned or second-planned, respectively), or was a new, entirely unplanned picture (see *Figures 3-5*). This third picture was presented for 2000 ms and was accompanied with a distinctive beep that alerted participants to the changed task requirements for these trials. The participants were told about these trials prior to the experiment and were instructed to name the picture as quickly as possible. Each trial started 500 ms after the end of the last, therefore, a trial lasted 5500 ms.

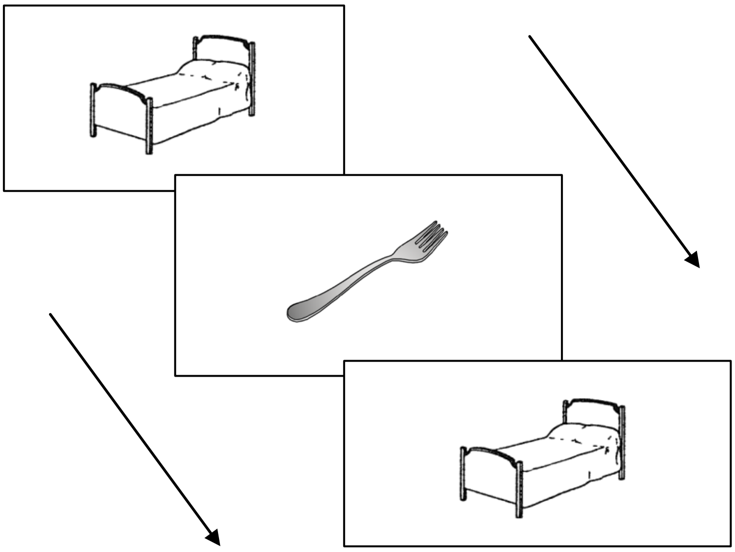

Figure 3. The catch trial procedure for the first-planned condition.

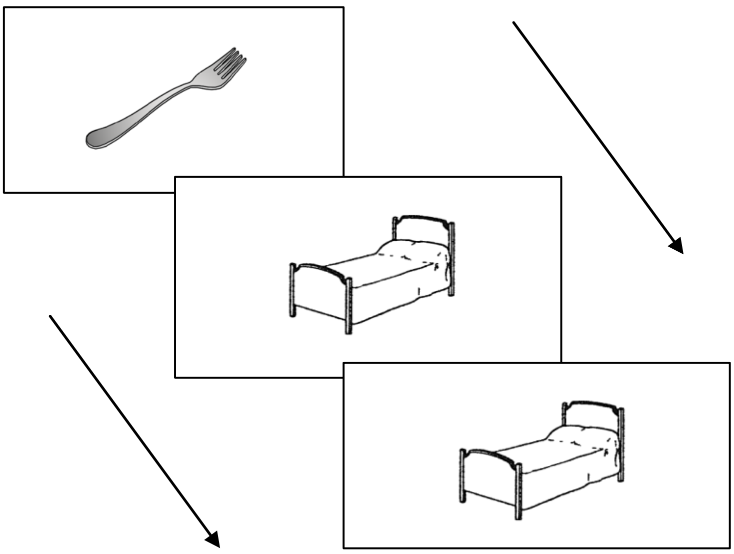

Figure 4. The catch trial procedure for the second-planned condition.

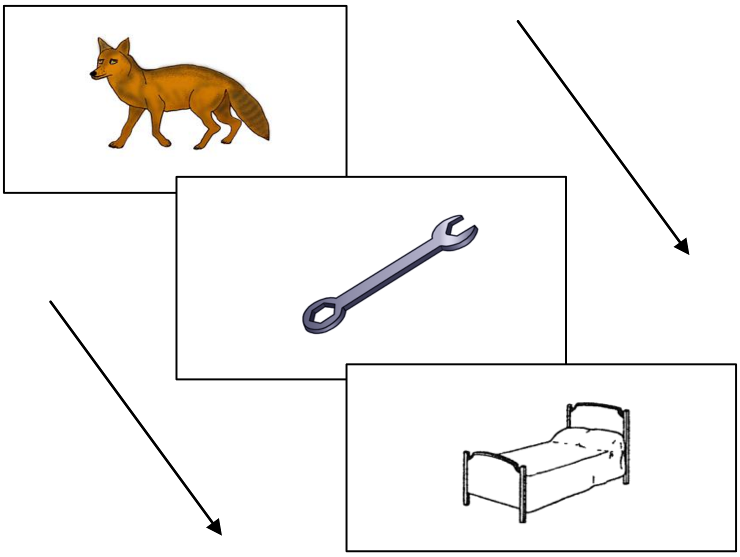

Figure 5. The catch trial procedure for the unplanned picture condition.

For each of the three experimental sessions, the same 36 picture pairs were used in the catch trials; what differed across sessions was the nature of the catch picture that appeared with each pair: whether it depicted the first-planned target, the second-planned target or an unplanned target. The 144 filler and 36 catch trials for each session were interspersed with each other and presented in a different, pseudo-random order across each of the three experimental sessions; sessions were balanced so that an equal number of different types of catch trials (first-planned, second-planned or unplanned) appeared in a single session (a full list of catch trial pictures and the order in which they appeared for each experimental session is given in Tables 3-5 in the Appendix).

At the beginning of the very first session, the participant completed a consent form, was given an information sheet and then proceeded with the picture naming task. For the next three experimental sessions, each session began with six practice trials, the last of which was a catch trial. The participant was given a short break half-way through the session. Also, just before commencing the second half of the experiment, participants completed another four practice trials, the last of which was a catch. The pictures used in the practice trials were not used in the actual experiment. Participants were instructed to watch the pictures being presented and to remember the order in which they appeared; after this, the pictures would then have to be said aloud at a prompt in the presented order with the word 'and' in between. They were also told that on a small number of trials, either the first presented picture, the second presented picture or a third, previously unseen picture would appear instead of the prompt. In this situation, they were told to name the picture as quickly as possible without compromising accuracy. A typical session lasted from around 20 to 25 minutes depending on the time it took for the instructions to be said and how long the participant took for a break. Participants were given a debriefing sheet at the end of the last session, as well as a verbal debriefing. Only one person took the experiment at a time.

Data and statistical analysis

Each session was recorded on a mini-disc player and was later manually analysed on the audio editing program Audacity (Mazzoni, 2000). As mentioned above, when the target picture appeared in a catch trial, there was a simultaneous, distinctive beeping noise. The onset latency measured was the time between this beep and the onset of the participant's response. Because each phrase appeared in three different catch-trial conditions (first-planned, second-planned or unplanned), there were three measures for each target word. Filler trials were not analysed in any way.

Prior to statistical analysis, the latency data for the catch trials was prepared as follows. First, any catch trials that were not correctly named were removed. Second, if the same target word was missed in at least two out of three sessions in which it was presented, that word was removed entirely from the participant's data. Third, outliers were removed, which were defined as latencies under and above two standard deviations of the grand mean for each participant. The resultant dataset was submitted to a General Linear Mixed Model analysis (or "mixed effects" model): the model incorporated two random effects - participant name and picture name.

Results

Accuracy analysis

A total of 7.7% of all catch trials were eliminated from the data set. Of these 83 trials, 35 were errors, either substitutions of words other than the target picture name (e.g. cup - “mug”; branch - “tree”; 25 trials in total) or failures to respond altogether (10 trials in total). 14 of the 35 errors occurred in the unplanned condition, 12 in the first-planned and 9 in the second-planned condition. As far as the catch target words are concerned, errors were not distributed equally amongst them. For example, the target words 'Cup' and 'Pool' accounted for 15 and 9 errors respectively. Given the relatively low numbers of errors, no further analysis was carried out on the error data.

A further 48 individual catch trials were eliminated from the data set because they either yielded naming latencies that were either below or above two STDs of the grand mean (43 trials) or were trials where two of the three target words (two of three conditions) were missing (five trials). The data were generally normally distributed and so were not subject to log transformation before analysis.

Latency analysis

Figure 6. shows the mean naming latencies for all three conditions. The mean RTs for the first and second-planned and unplanned conditions were 604.33ms, 595.46ms and 742.07ms, respectively. The data were analysed in two different ways. In the first analysis, the predictor variables included catch trial type (first and second-planned and unplanned conditions) and session number (session1, 2 or 3), and the interaction between catch trial type and session. In the second analysis, session was not included as a predictor variable.

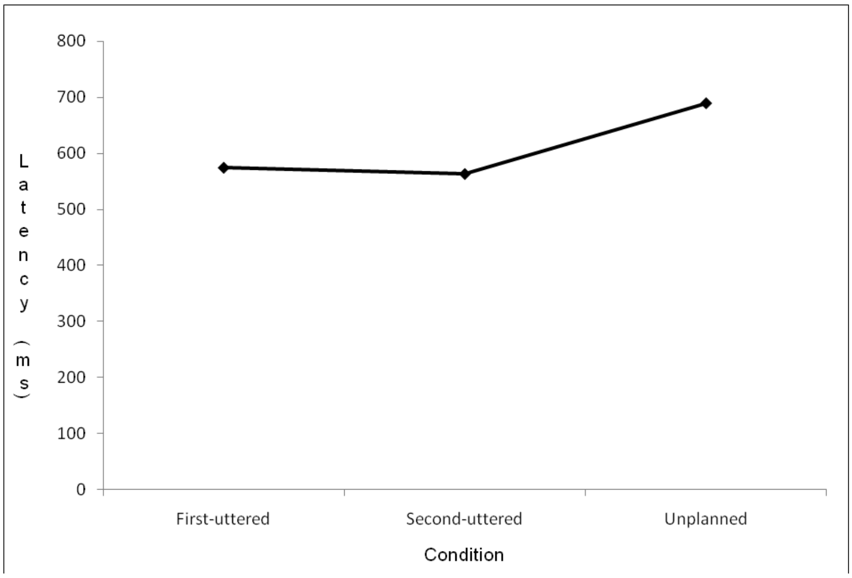

Figure 6. The mean RT latencies for all participants across all three conditions.

There was a significant main effect of session: participants became faster with each succeeding session, $F(1,947)=154.86$, $p<0.01$. There was no significant main effect of condition, $F(2,947)=0.64$, $p=0.5265$, but there was a significant interaction between session and condition, $F(2,947)=7.74$, $p<0.01$.

Second analysis

In order to further explore the effect of condition irrespective of session, a second analysis was performed which did not include the session by condition interaction. Running the analysis reveals a significant main effect of condition, $F(2,949)=145.29$, $p<0.01$ and session, $F(1,949)=154.75$, $p<0.01$. As can be seen in Figure 6, there is a clear difference between the two planned conditions (first-planned and second-planned) and the unplanned condition. Both the first-planned and second-planned conditions were significantly faster than the unplanned, $F(1,949)=197.40$, $p<0.01$ and $F(1,949)=242.49$, $p<0.01$, respectively. Although numerically, latencies in the first-planned condition were longer than those for the second-planned condition, this difference did not reach significance, $F(1,949)=2.39$, $p=0.1223$.

Discussion

Our prediction that naming responses in the first-planned condition would be faster than in the second-uttered condition was not supported. A faster first-planned condition was hypothesised because, being the first word to be spoken in the phrase, the corresponding lexicon should have the highest activation at the time just prior to production. Not only was there no significant difference between the first and second-planned conditions in this experiment, contrary to expectation there was actually a trend for higher latencies in the first-planned than in the second-planned condition.

Naming latencies in the unplanned condition were significantly slower than both the first-planned and second-planned conditions. It was hypothesised that if this condition produced RTs no shorter than those of the second-planned, then the ‘boosting’ hypothesis would be supported (Dell, 1986; Freedman et al., 2004). However, if the unplanned condition produced RTs that were shorter than those of the second condition, then the ‘inhibition’ theories would be supported. The present findings may appear to support activation-boost theories - that is, those that suggest items planned for upcoming positions in an utterance are not inhibited, but rather receive an additional boost to their activation levels just prior to their production, but the first-planned condition RTs were not shorter than those of the second-planned condition, as per our first hypothesis. Therefore an alternative conclusion must be considered. Latencies in the unplanned condition were actually extremely delayed above the other two context conditions. At least some of this delay may be attributable to processes occurring prior to lexical access and selection, such as picture identification and/or semantic access. Recall that the unplanned picture was not exposed prior to the onset of its presentation as the catch picture. Unlike the first and second-uttered words, whose corresponding pictures had already been viewed just prior to the catch, these unplanned pictures enjoyed no such benefit. Therefore, it is most likely that the exaggerated delay seen in the unplanned condition is due to the lack of exposure of the pictures in this condition prior to phrase production.

A second potential problem with Experiment 1 is that the effect of condition was powerfully modulated by session number. This raises the possibility that some of the condition effect may be attributable to strategic factors. For example, since different sets of pictures were used in filler trials and catch trials, and those used in catch trials were used several times over, it is possible that over the course of the testing sessions, participants learned which pairs of pictures were likely to be followed by a catch trial. For these types of picture pairs participants may not have planned to produce the target phrase as enthusiastically as they did on the filler trials.

Experiment 2

The results from Experiment 1 indicated two critical design issues which Experiment 2 was designed to address. The first issue was that the pictures that appeared in the unplanned condition were not given enough exposure prior to production of the picture name. The second issue was that participants became much quicker at naming the catch trial pictures as a function of session number, which meant that they were remembering the catch trial pictures and hence when a catch trial was forthcoming.

In Experiment 2, participants completed the same primary task as in Experiment 1: that is, they had to produce a noun phrase (two nouns conjoined with "and") in response to a picture prompt. However, rather than viewing two pictures successively, participants viewed an array of three pictures, presented simultaneously. The phrase to be produced was indicated through the use of animation: First, the target to be produced in first position in the phrase was animated (the picture bounced up and down on the screen; henceforth known as first-planned condition). Then following this, the target to be produced in second position was animated (this picture flashed; henceforth known as second-planned condition). The third picture, not to be included in the target phrase, remained stationary (henceforth known as unplanned condition). The spatial position of the pictures was varied, so its role in the target phrase could not be predicted from its location but could only be learned from viewing the animation sequence. This setup was designed in order to address our first methodological concern in the previous experiment. During the unplanned condition in Experiment 1, the picture name to be produced was not seen and therefore planned just prior to production. In this experiment, the picture which was required to be produced in the unplanned condition was shown just prior *as* the stationary picture. The participants saw three pictures, two of which were animated and were told that they were to name the animated pictures only. In this way, the participants were exposed to the unplanned picture whilst not actually planning to produce it later in the trial.

The second methodological concerned raised in Experiment 1 was the participants' decrease in RT for the catch trials across the experimental sessions (remember, there were three experimental sessions). In order to address this problem, the participant pool was divided into three and the picture lists for the three experimental sessions were rotated amongst them in a counter-balanced order. So, one third of the pool started on one picture list (which we shall arbitrarily name 'list A'), then next session were shown the next list of catch trial pictures (list B) and were finally shown the last list for the final session (list C). Another third of participants was shown list B for the first session, followed by list C for the second and finally list A. The final third of participants were shown list C first, followed by list A, then list B. In this way, any RT advantages that were produced as a result of learning would cancel out when all participants' data were pooled.

One further variable that was examined in this experiment was the semantic relationship between the two items planned for the target phrase. In the studies conducted by Freedman et al. (2004) and Scott and Wilshire (2008), latencies for pictures presented in a semantically competitive context increased compared to a non-competitive context. Manipulating the semantic context allows us to examine phrase production in more detail. Even though Freedman et al. and Scott and Wilshire advocate different mechanisms for how semantic competition is resolved, both parties argue that such contexts tax the lexical selection process more so than non-competitive contexts. In the present study, latencies for words in different places within a phrase are being evaluated and it would be of interest to see whether semantic competitive contexts tax items across all three conditions in an equal fashion, or if differences in RTs will arise due to the place in which the item is planned within the phrase.

The predictions are the same as for Experiment 1; that is naming latencies are predicted to be faster when the picture depicts the word planned for first position in the original phrase (first-planned condition), than when it depicts the word planned for second position (second-planned condition). This is because the lexical unit for this first-planned target, being the first item in the phrase to be produced, should already be activated above its second-planned competitor at the time the catch-trial target is presented. Regarding the unplanned trials, if naming latencies for unplanned catch trials are no quicker than those of second-planned catch trials, then "activation boost" theories will be supported (e.g., Dell, 1986; Freedman et al., 2004). This is because the first-planned word is receiving the boost and the unplanned word is not engaged in any way during planning. Conversely, if unplanned condition latencies are faster than those of second-planned catch trials, then "inhibition" theories are supported (Wilshire and McCarthy, 2002; Hamilton and Martin, 2005; Biegler et al. 2008). This is because the lexical unit for the word planned for the second position of the phrase is being inhibited, whereas the unplanned picture is not.

Regarding semantic contextual manipulation, it is hypothesised that target pictures in catch trials that have just been simultaneously presented with a semantically related picture will take longer to name than those presented with unrelated pictures. The reason for this is because competitive contexts tax the purported selection mechanism more, thereby causing a delay in word selection (Wilshire and McCarthy, 2002; Hamilton and Martin, 2005; Biegler et al. 2008; Scott and Wilshire, 2008). The competition stems from the fact that semantically related items share particular features - for example, cats and dogs are both furry have whiskers, are often pets etc. This overlap in semantic features taxes the selection mechanism more because it must select which lexical unit is the best choice, given the particular activated semantic units.

It is hard to predict what varying the semantic relatedness will have on the relative timings of the catch words; however, if an interaction between noun utterance position and semantic relatedness is found, then exploring it might reveal more information about sentence processing.

Method

Participants

Participants were 24 first year psychology, native English-speaking students studying at Victoria University of mixed sex, although most of them were female (n=17). 31 participants were signed up for the experiment but seven participants' data were eliminated because the participants were not focused enough on the tasks and made too many mistakes.

Materials

The picture stimuli for this experiment consisted of 551 line drawings of typical objects and animals, belonging to one of seven semantic categories: people, animals, fruit, vegetables, furniture, utensils and tools. The drawings were taken from a variety of sources, consisting of: photos, black line drawings and coloured drawings. The pictures were sorted into two groups: those that were used in filler trials (357) and those that were used in catch trials (194; see below in '*Design and procedure*' for details about filler and catch trials). 408 of these pictures were used in Experiment 1: 302 of these pictures were used in both experiments as filler trial pictures and 106 pictures were used in both experiments as catch trial pictures. The filler trial pictures had varying syllabic lengths. The catch trial pictures had monosyllabic names. The Kucera-Francis (1967) written frequencies for the catch trial pictures ranged 1-1207 per million with a mean of 73.79 and a standard deviation of 156.58 (Wilson, 1987).

The 357 filler trial pictures were used to create 144 different picture triplets – this was done three times to create three different sets of triplets for each of the three experimental sessions. In half of these triplets, two of the three pictures were semantically related (e.g. Kilt, Jersey, Pill); for the other half, all three were unrelated (e.g. Knee, Razor, Ship). There were not enough pictures for every filler trial within a session to have different pictures, so some pictures had to be used twice; however, filler pictures that appeared twice in one session were always shown in different triplets.

The 194 catch trial pictures were used to create 216 triplets - or to be more precise, 6 sets of 36 triplets. Only one set of 36 triplets was used in a session (which, plus the 144 filler trials, gives us our 180 trials per experimental session). However, different sets of triplets were used in different sessions (more on this later). Of these 194 catch trial pictures, 36 of them were the targets, which were the pictures that were shown again instead of the response cue. The remaining 158 catch trial pictures were used as the other pictures in the triplet (a full list of catch trial pictures for Experiment 2 can be found in Table 6 in the Appendix). As with the filler trials, there were two types of catch trial triplet: one with two semantically related items in it and a triplet with no related items in it. When creating the triplets, a target picture would either be paired up with a semantically related item as well as an unrelated item *or* with two unrelated items. A further constraint was that the target pictures in a given triplet did not start with the same letter.

Each of the 36 target pictures had to appear in each of the three plan conditions; therefore, each target picture was put into three different triplets: one with it as the first-planned word (bouncing), the second-planned word (flashing) and the unplanned word (stationary). These three different triplets for each target picture were used across the three sessions so that each target appeared in a different condition each session. Different accompanying pictures were used in these different triplets, whether they were semantically related or not. It must be noted that in any particular session, the 36 target pictures were always separated into three subgroups of 12: one for each plan condition. Therefore, in any session, 12 pictures would be the first-planned word (bounce), 12 would be the second-planned word (flash) and 12 would be the unplanned word (stationary).

As just outlined, half of the catch trial triplets had two pictures that were semantically related and half did not. Another set of triplets were created that had the target pictures which had been semantically matched up originally, now placed into triplets with items that were not semantically related. Conversely, the target pictures that had originally been placed in unrelated triplets were placed into triplets that had another picture that was related. This resulted in six sets of triplets, which allowed every target picture to appear in all three plan conditions (first-planned, second-planned, unplanned), in each of the two semantic conditions (related, unrelated).

There were six sets of catch trial pictures – in one group of three, half the target pictures were semantically related to another item in the triplet and half were not. In the other group of three, the pictures were swapped so that the unrelated target pictures were now placed with related pictures. Apart from this semantic change, everything was kept the same between the two groups (the presentation order, the other catch trial pictures in the triplet, the order of the filler trials etc.; see Tables 7-12 in the Appendix for a list of all the catch trial pictures for each experimental session, for each group). The 36 catch trial triplets were added to 144 filler trial triplets. As there were three sets of 36 catch trials (per group: remember there were two groups), the filler trial triplets were remixed three times. Once reshuffled, the filler triplets were added to the appropriate set of 36 catch trial triplets and mixed together. This resulted in three full sets of experimental sessions with 180 trials. Once the target pictures had been re-paired with either a semantically related or unrelated item (depending on how it was paired originally; discussed above), there were six full sets.

The pictures were shown on a Macintosh computer and each session was taped using a mini-disc recorder.

Design and Procedure

Each participant was required to complete four sessions. The first session involved a standard naming task that was identical procedurally to that used in Experiment 1, but with a different pool of pictures. The remaining three sessions comprised the actual experiment. As with Experiment 1, each session consisted of 180 trials: 144 of these were filler trials and the other 36 were catch trials. For the filler trials, the procedure was always the same: first, the three pictures in the stimulus triplet were displayed simultaneously from left to right across the screen (see *Figure 7*). 200ms later, the picture target to be produced first was animated (it was displaced directly above its original location for 50 ms, then displaced even higher for another 50 ms, then this animation sequence was repeated a second time, producing a bouncing appearance). Then 800 ms later, the picture target to be produced second was animated (this picture disappeared for 50 ms, then re-appeared for 50 ms, then disappeared and reappeared again for a further 50 ms each time). The third picture target, which was not destined for the target phrase, remained stationary throughout. 1400 ms after the completion of the animation sequence, a blank screen appeared, followed 1500 ms later by a response cue ("____and____"). At this time, the participants were required to produce the target phrase. One complete trial lasted 8000 ms. The physical position of the first-planned, second-planned and unplanned target pictures was systematically varied so that a picture's role in the target phrase could not be deduced by its physical location .

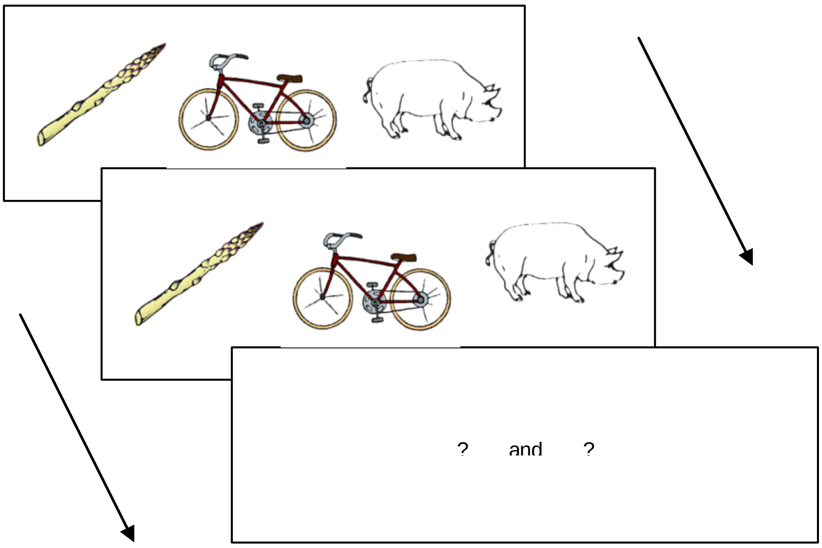

Figure 7. The procedure for the filler trials

In the 36 catch trials, the procedure was exactly the same as that mentioned above up until the prompt ("____and____"). Instead of seeing the prompt, the participants were shown one of the three pictures that had just previously been presented: the first-planned, the second-planned or the unplanned picture target (see

Figures 8-10). This picture was presented for 2000 ms and its onset was accompanied by a distinctive beep. The participants were warned of these trials and were instructed to name the single picture as quickly as possible. The filler and catch trials were randomly mixed before the experiment for all three sessions and were presented in a fixed, random order. Unlike in Experiment 1, catch trials never appeared twice in a row: there was always at least one filler trial in between them.

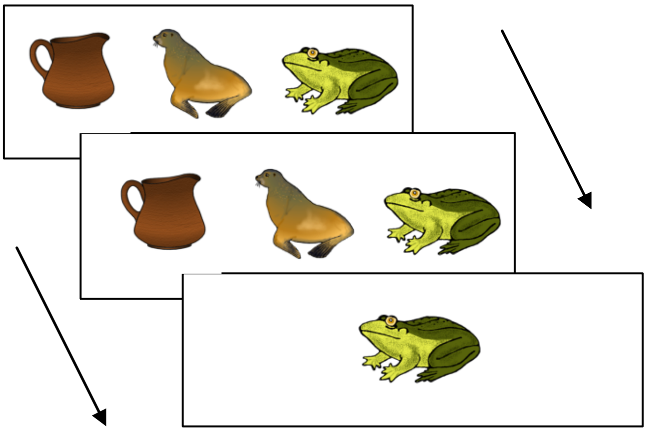

Figure 8. The catch trial procedure for the first-uttered condition (in this picture, the frog is bouncing and the seal is flashing).

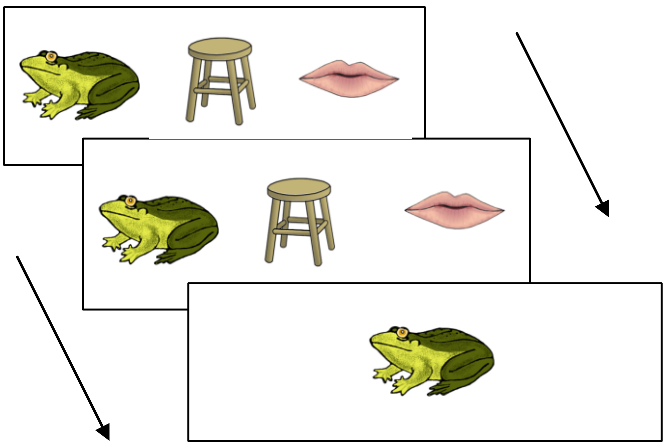

Figure 9. The catch trial procedure for the second-uttered condition (in this picture, the stool is bouncing and the frog is flashing).

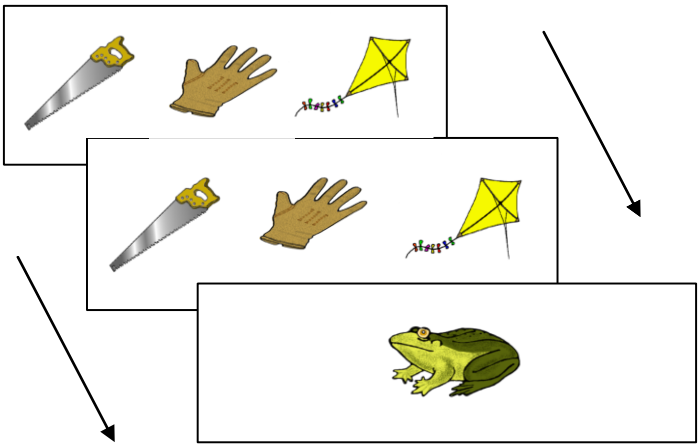

Figure 10. The catch trial procedure for the unplanned condition (in this picture, the glove is bouncing and the kite is flashing).

Throughout the entire experiment, the same set of 36 target pictures was used as targets for the catch trials (they appeared again instead of the cue; ____?____). Much like in Experiment 1, the set of 36 pictures was split into three subgroups of 12. In one session, one group of 12 pictures was the first-planned condition pictures that were meant to be said first but which instead appeared again instead of the cue. During this same session, another group of 12 pictures was the second-planned pictures. The last group of 12 acted as the unplanned condition pictures. For example, if the target picture in the catch trial was a cat, then in one session, it would be in the first-condition and would therefore bounce, then a picture of, say, a desk would flash while a picture of a cup remained stationary. Then where the cue would normally appear, the picture of the cat would appear again instead. When the participant would come back for the next session, the picture of a cat would come up again (remember, the 36 target pictures were shown in every session), but it would be in another condition. So, if in this session, the cat was appearing in the second-planned condition, the participant would see a van bounce, then the cat flash while a tree remained stationary. The cat would then appear again instead of the cue. In the last session (in which the cat would appear in the unplanned condition), the participant may see a fork bouncing, followed by a screw flashing while the picture of a cat remained stationary. Again, the cat would appear afterwards instead of the cue (see Table 1 below; a full list of catch trial pictures and the order in which they appeared for each experimental session is given in Tables 7-12 in the Appendix).

Table 1

An example of three catch trial triplets with the same target picture (cat) used across three experimental sessions.

Session	Condition First-planned (bounce)	Second planned (flash)	Unplanned (stationary)
1	Cat	Desk	Cup
2	Van	Cat	Tree
3	Fork	Screw	Cat

In order to counter-balance the 36 target pictures, the three subgroups of 12 were arbitrarily given a set letter: A, B or C. In Experiment 1 there had been a degree of learning regarding the repeated catch trial targets, which affected the onset latencies. Therefore, in Experiment 2, a third of the participants started with set A as the first-planned condition pictures, set B as the second-planned and set C as the unplanned pictures. Another third would have set B as the first-planned pictures, set C as the second-planned pictures and set A as the unplanned pictures. The last third of the participants had set C as the first-planned pictures, set A as the second-planned pictures and set B as the unplanned pictures. By counter-balancing in this way, any RT differences for the word position condition for each picture that appeared as a function of session canceled out.

The placement of the target picture was also counter-balanced. Recall that the there were three places across the screen where the pictures simultaneously appeared. In one session, a target picture (e.g. cat) would appear in the middle position. In the next session, the picture would appear in the left position and in the last session, it would appear in the right position. Therefore, each target picture made all three actions and appeared in all positions across the three sessions.

To evaluate the semantic relatedness effect, the entire group of participants was split into two groups. In one group, half the catch trial target pictures were preceded by a semantically related picture and the other half was not. For the related half, when the catch was the first-planned picture (bounce), the second-planned picture (flash) was that which was semantically related; when the catch picture was the second-planned picture, the first-planned picture was semantically related; when the catch was the unplanned picture, the first-planned picture was semantically related. The other group had exactly the same order and format but the pictures that were preceded by related objects were now preceded by unrelated ones and vice-versa for the other half of the targets.

For the first session, the participant completed a consent form, was handed an information sheet and then underwent the picture naming task. For the next three experimental sessions, the procedure was as follows: an experimental session was split into two halves, in order to give the participant a short break half-way through. Before the first half, participants completed a practice run of three trials, the last of which only was a catch. Also, just before the second half of the experiment, participants completed another practice run which also consisted of three trials; again, only the last trial was a catch. The pictures used in the practice run were not used in the actual experiment. Participants were instructed to watch the pictures being presented and to remember which one bounced and which one flashed; after this, the pictures would then have to be said aloud at a prompt, at which the bouncing picture would have to said first, followed by the flashing one with the word 'and' in between. A typical session lasted from around 25 to 30 minutes depending on the time it took for the instructions to be said and how long the participant took for a break. Participants were given a debriefing sheet at the end of the last session, as well as a verbal debriefing. Only one person took the experiment at a time.

Data and statistical analysis

Each session was recorded on a mini-disc player: response latencies were then obtained manually using the audio editing program Audacity (Mazzoni, 2000). As mentioned above, when the target picture appeared in a catch trial, there was a simultaneous beeping noise. The onset latency measured was the time between this beep and the onset of the participants' response. Filler trials were not analysed in any way. Prior to statistical analysis, the incorrect responses were removed from the data. Outliers, defined as latencies under and above two STDs of the grand mean for each participant were also removed. Finally, if two of the three targets for any word were incorrectly responded to, then that whole word was removed from the data for that participant. When the latency data was collected, each target word for each participant had three RTs, one for each condition, not including errors. Data for the participants were submitted to a General Linear Mixed Model analysis (or "mixed effects" model): the model incorporated two random effects - participant name and target picture name.

The analysis examined the role of prior context on the latency of response for each catch trial.

Results

Accuracy analysis

A total of 5.29% of all catch trials were eliminated from the data set. Of these 137 trials, 56 were errors, of which the majority was word substitutions ((e.g. bee - “wasp”; frog - “toad”; gun - “pistol”; mouse - “rat”; 50 instances); the remainder was omissions, where the participant failed to make any response whatsoever (6 instances). With respect to the three conditions in which a picture could appear, the most error prone was the second-planned condition. This condition accounted for 25 of the 56 errors (44.6%). The first-planned and unplanned conditions accounted for 12 (21.4%) and 19 (33.9%) errors respectively. As far as the catch target words are concerned, errors were not distributed equally amongst them. For example, the target words 'Bee' and 'Plane' accounted for 19 and 15 errors respectively; whereas, the target words 'Bed' and 'Pear' each only had one error. A further 81 catch trials were eliminated from the data set because they either yielded naming latencies that were either below or above two STDs of the grand mean or were trials where two of the three target words (two of three conditions) were missing. The data were generally normally distributed and so were not subject to log transformation before analysis. Given the relatively low numbers of errors, no further analysis was carried out on the error data.

Latency analysis

The data were generally normally distributed and so were not subject to log transformation before analysis. The mean RTs for the first-planned, second-planned and unplanned conditions were 635.99ms, 633.9ms and 698.71ms, respectively. Overall, both the first-planned and second-planned condition were significantly different from that of the unplanned, $F(1,2391)=89.64$, $p<0.01$ and $F(1,2391)=109.86$, $p<0.01$, respectively; but they themselves did not differ significantly, $F(1,2391)=1.12$, $p=0.2899$. *Figure 11* shows the mean naming latencies for all three conditions, as a function of semantic relatedness. As can be seen from Figure 11, there is a clear difference between the related and unrelated groups, which reached significance, $F(1,2391)=19.92$, $p<0.01$. The interaction between condition and relatedness did not reach significance $F(2,2389)=0.41$, $p=0.6643$.

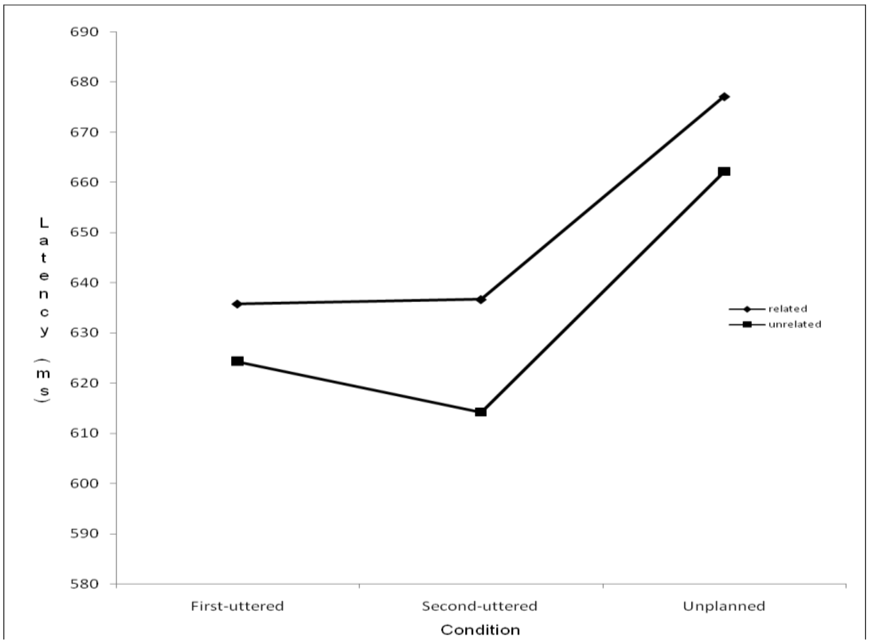

Figure 11. Mean naming latencies for all conditions.

Discussion

Despite our modifications to the experimental method, the findings in Experiment 2 were very similar to those of Experiment 1: naming latencies for the unplanned picture condition (for both the related and unrelated pictures) were significantly longer than those for both the first-planned and second-planned conditions, which themselves did not differ significantly. It appears as though our modifications did not substantially impact on the findings obtained. One possible explanation for this outcome is that the unplanned condition pictures may still not have been exposed enough prior to their verbal production. The unplanned pictures were shown with the other pictures during the animations (bouncing and flashing) but because the participants only had to remember the animated pictures, it may be that these pictures received much more processing than the unplanned pictures. For example, as soon as the animations were complete and the participants were waiting for the cue, they could have been picturing only the animated pictures in order recall them at the cue. This may have shadowed the retention for the unplanned picture.

In Experiment 2, we also manipulated the semantic relationship between the items planed for first and second position in the target phrase. In general, catch trials involving these semantically related triplets produced longer latencies than those for unrelated, as per our hypothesis. This was expected because the overlap of features at the semantic level would have introduced more competition at the lexical level; this result is in line with the earlier research of Freedman et al. (2004) and Scott and Wilshire (2008).   We were also interested in whether there would be an interaction between relatedness and word position. A prediction was not made because the search for an interaction was exploratory and not guided by a hypothesis. It was supposed that if an interaction was discovered, then evaluation of it may have shed more light on the process of speech production in competitive contexts. However, no interaction between these two variables was found.

General Discussion

The two experiments described here set out to examine how exactly a word's lexical unit was selected amongst others in order to be produced verbally as part of a short phrase. We encouraged participants to plan simple two-noun phrases that they subsequently had to produce. We then endeavored to probe this process by presenting catch trials 20% of the time, in which instead of producing their planned phrase, they were re-presented with one of three target pictures. The unplanned picture condition was intended to act as a baseline to compare the onset latencies for the first-planned and second-planned pictures. Experiment 2 was also set up to examine the lexical selection process; however, the unplanned picture was shown simultaneously with the other two in order to facilitate a deeper processing of it. In both experiments, it was hypothesised that the RTs for catch trials involving the first-planned word would be faster than those for the second-planned word. This is consistent with the majority of theories (see *Theories of single word production* and *Production in context* above), which propose that just prior to initial of a multi-word utterance, the items to be produced first will be more highly activated than those to be produced later in the same utterance. In both experiments, the naming latencies for the first-planned and second-planned words were not significantly different from each other. In fact there was a very slight trend for the first-planned word to have a longer naming latency than that of the second. However, this was not significant in either of the current studies. It may be premature to conclude anything as far as the relative timing for the first and second words are concerned, as it could be that such an RT difference, if it did exist, would need a more sensitive methodology to expose it.

Another possibility of why a non-significant finding regarding the relative timings of the first and second-planned words resulted may be because the process of uttering the phrase is one that is cognitively distinct from that of lexical planning: it may be that the phonological process is one that is not getting 'caught out' by the catch trials. Presenting the catch picture would access lexical representations (from the semantic features), just as presenting the two nouns earlier in the trial was. However, when the participant is just about to utter the phrase, it may be that catching the participant out by accessing lexical units is not appropriate because the phonological process may be too far removed from that of the lexical and/or it may be a process that is much quicker than lexical retrieval. In Dell's spreading activation model, after a lexical unit has been selected, activation is feed into the appropriate phonological units so that the item can be verbalized (Dell, 1986; see figure 1). In the present paradigm, when a participant is 'caught out', they are presented with an unexpected picture. This would feed in through the semantic level, then proceed to that of the lexical and finally to the phonological level. It may be that that attempting to catch the participant out this way is not appropriate because it takes too long to catch the phonological processes out by probing the lexicon from the semantic level. It may be more fruitful to catch the phonological processes out by interrupting them with a phonological probe instead. An alternative possibility for future research may be to introduce another modality into the paradigm; this idea is discussed below in *'Proposals for future research'*.

Regarding the catch trials involving unplanned words, no hypothesis was made. However, it was determined that if naming latencies for the unplanned catch trials were no quicker than those of second-planned catch trials, then "activation boost" theories would be supported (Dell, 1986; Freedman et al., 2004) because the first-planned word is receiving the boost and the unplanned word is not engaged in any way during planning. However, if unplanned condition latencies were faster than those of the second-planned catch trials, then "inhibition" theories would be supported (Wilshire and McCarthy, 2002; Hamilton and Martin, 2005; Biegler et al. 2008) because the lexical unit for the word planned for the second position of the phrase is being inhibited, whereas the unplanned picture is not. Neither theory is supported by the present findings because the unplanned catch trial latencies were greatly exaggerated above those of the first and second-planned conditions. Although this may seem to support the activation boost theories, the fact that there was no significant difference between the first and second-planned conditions means that we cannot conclude this, as it was also hypothesised that the first-planned condition would have shorter RTs than the second-planned condition, if its activation is indeed getting a boost.

These results may reflect the difficulty inherent in formulating an appropriate "unplanned" baseline. Even in Experiment 2, where the unplanned item was viewed as part of the initial target triplet, it may not have been attended to as much as the other two. So, as discussed above, there may still be differences between the planned and unplanned conditions that reflect processes outside lexical activation and selection

Experiment 2 also increased the competitive context by manipulating the semantic relationship between the two words planned for the phrase; this manipulation was exploratory. A significant delay was found when catch trial target pictures were related to one of the pictures presented with it seconds earlier in the trial; a finding in line with previous research (Freedman et al., 2004; Scott and Wilshire, 2008). Although such a finding supports previous research, it does not help to delineate which purported theoretical model best describes the lexical selection process because such a finding can be described by the boosting model or the inhibiting model.

There was a non-significant numerical trend suggesting that the second-planned condition RT was shorter than that of the first in Experiment 1 and for the unrelated group in Experiment 2 (which was much like Experiment 1 because there was no semantic manipulation in Experiment 1). This trend disappeared in Experiment 2 for the related group. If this trend were found to be statistically reliable in future studies, it might suggest that there is less of an activation differential between the two planned words when they are semantically related than when they are not. This converges with other evidence to suggest that there is greater competition between semantically related words planned for the same phrase than for unrelated words (Freedman et al., 2004; Scott and Wilshire, 2008). In fact, this is another way of demonstrating the effect in a task that does not explicitly require production of the planned phrase in its entirety.

Proposals for future research

For future research employing the catch procedure, it would be more fruitful to investigate the relative timings for words to be uttered, rather than continue with using the unplanned picture as a frame of reference. It is interesting that there was a slight trend towards the first word's RT being longer than that of the second across the two experiments. It would be of theoretical interest to ascertain the facilitation of a word as a function of its position in a sentence. In an effort to understand more about this facilitation, future research would perhaps benefit from more words being used in the sentence; for example, simply including another word (e.g. dog and tree and arm) and measuring the differences in RT between all the words may help expose facilitation differences, should they exist. This is because the sentence is a bit longer and therefore more opportunity is afforded to interrupt planning processes. If differences of facilitation between words of different utterance position do exist, then it would be easier to discover them by looking at the difference between words that have a degree of space between them. For example, we could look at the difference between the first and third word in an utterance (dog and arm in the above example). Also, three words increases what we may be able to interpret from the findings; for example, it introduces the possibility of establishing linearity, should the RTs increase or decrease as a function of utterance position.

It would also be interesting to discover what may happen when the catch trial pictures also include pictures of items that are semantically related to the pictures of the words that are being planned for an utterance and the effect this would have on the relative timing of the words of the utterance. Future research may want to set up the paradigm employed by this study, where the participant witnesses two (or three) unrelated pictures which are to be named. However, as well as catching the participants out with either of the just-presented pictures or some control picture, semantically related pictures could be used instead and the baseline could be dropped qua a baseline (as far as this term has been used in this study) and be used as a simple control instead. For example, if two pictures are going to be shown, say, a cat followed by a shirt, then the participants would normally be expected to utter “cat and shirt” at a given prompt. However, on catch trials, instead of the prompt, a picture of a cat or a shirt *or* a pig or a hat could be presented instead, following which, the participants would be required to name them aloud as quickly as possible without compromising accuracy. Catch trials would also contain random pictures that were not semantically related to either of the previously presented pictures (e.g. an arm), which would act as a control against which the just seen or semantically related pictures could be measured.

In such an experiment, we would expect (based on the findings of the current study) that the trials with the random picture shown as a catch trial to have the highest RTs. If the first, second or third picture of the utterance were shown again as a catch, it is hard to predict which would have the shortest RT. Again, it is hypothesised that it would be the first word, and it is presumed that the reason why such a finding was not discovered in the present study was due to methodological shortcomings. Again, this prediction is made because the first word should be on the 'tip of the tongue'. It is also hypothesised that the last word of the utterance should have the longest RT, as it is the last word planned. Regarding the semantically related pictures, again it is hard to predict (based on current findings) what may happen. In experiment 2, planning two words that were semantically related and catching with one of them caused a delay in onset latency. However, in the proposed study, the planned words are semantically unrelated: it is the catch that is related. As the word is not planned, it will have a longer RT than those of the repeated catch pictures (where the catch picture is one of those just seen). However, due to the semantic relatedness, there could in fact be a priming effect which would lead to shorter RTs than those of the random pictures. The interesting issue would be the interaction between the semantic variable and the position variable. If the semantic variable resulted in RTs different from those of the random pictures, then it would be important to see if they also differed as a function of utterance position. Above, it was hypothesised that the last word of the utterance would have the longest RT. If the picture shown was semantically related to the last word of the utterance, and its RT was slower that those of pictures related to the first and second word, this would mean that the planning processes affected semantic neighbours as well as words to be uttered.

It was mentioned above that the participant may not have been caught out when preparing to utter the two-noun phrase, because the process which we were hoping to interrupt was phonological, whereas the catch manipulation was one which relied on semantic units feeding into lexical units, which then feed into phonological units. By the time the phonology for the catch picture would have been processed, it may have been too late to interrupt the planed noun phrase. A better way of interrupting such a process may be to catch the participant out with a verbal item instead. That is, have the participant view the pictures and prepare to utter the phrase at the prompt, but instead have a recording of an item's name play aloud (instead of showing another picture), which the participant must utter. The benefit of this is that the verbal presentation of the word is closer to the phonology than that of the visual. Upon seeing a picture, semantic units activate lexical units, which then activate phonological units if the picture is to be named. However, hearing a word works the other way round; that is, upon hearing a word, the phonological units activate the lexical units, which in turn activate the semantic. Therefore if, upon seeing two pictures, the participant eventually engages in phonological processing just before uttering the phrase, interrupting them with phonological information may be a more relevant way to catch the participant out. Indeed, there is already evidence in the literature that such an effect exists within a paradigm that is called the auditory picture-word interference task (Wilshire et al., 2007). In such a paradigm, participants must name a picture whilst ignoring an auditory item. Typically, there is an increase in picture naming when the auditory item is semantically related to the picture. Of course, these suggestions need not be taken as separate proposals; the idea of introducing one more word to the utterance, using semantically related catches and introducing verbal catches may be incorporated, producing an experiment that may produce results in line with the predictions made before the current experiments were conducted.

The above proposals would be theoretically interesting because it may help ascertain what is happening with the elements of a sentence when we are planning to speak. The baseline idea was discouraged above, but that does not mean that there may not be a way to uncover whether the boosting of inhibiting hypotheses best describes sentence processing. The above proposal may give us a partial answer to this question. For example, if the catch picture used was semantically related to the last word to be spoken in the utterance, and there was an interaction between the position and the semantic relation such that the RT that resulted was high because of the position *and* lower still because of the semantic relation, then that would lend credence to the inhibition hypothesis. This is because such a finding would lend support to the hypothesis that the selection mechanism has its effect through inhibition.

References

Badecker, W., Miozzo, M., & Zanuttini, R. (1995). The two-stage model of lexical retrieval: Evidence from a case of anomia with selective preservation of grammatical gender. *Cognition, 57,* 193-216.

Barde, L.H.F., Schwartz, M.F., Chrysikou, E.G., & Thompson-Schill, S.L. (2010). Reduced short-term memory span in aphasia and susceptibility to interference: Contribution of material-specific maintenance deficits. *Neuropsychologia, 48*, 909–920

Biegler, K.A., Crowther, J.E., & Martin, R.C. (2008). Consequences of an inhibition deficit for word production and comprehension: Evidence from the semantic blocking paradigm. *Cognitive Neuropsychology, 25 (4),* 493-527.

Dell, G.S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93 (3),* 283-321.

Dell, G.S., & O'Seaghdha, P.G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review,98 (4),* 604-614.

Freedman, M.L., Martin, R.C., & Biegler, K. (2004). Semantic relatedness effects in conjoined noun phrase production: Implications for the role of short-term memory. *Cognitive Neuropsychology, 21 (2/3/4),* 245-265.

Hamilton, A.C., & Martin, R.C. (2005). Dissociations among tasks involving inhibition: A single case study. *Cognitive, Affective and Behavioural Neuroscience, 5 (1),* 1-13.

Hamilton, A.C., & Martin, R.C. (2007). Proactive interference in a semantic short-term memory deficit: role of semantic and phonological relatedness. *Cortex, 43,* 112-123.

Kucera and Francis, W.N. (1967). Computational Analysis of Present-Day American English. Providence: Brown University Press.

Mackay, D.G. (1987). *The organization of perception and action.* New York: Springer Verlag.

May, C.P., Hasher, L., & Kane, M.J. (1999). The role of interference in memory span. *Memory & Cognition, 27 (5),* 759-767.

Mazzoni, D. (2000). Audacity (Version 1.3.12) (Software). Santa Monica, CA.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1-86.

McClelland, J.L., & Rumelhart, D.E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of the basic findings. *Psychological Review, 88,* 375-407

Rapp, B., & Goldrick, M. (2000). Discreteness and Interactivity in Spoken Word Production. *Psychological Review,107,* 460-499.

Schwartz, M.F., Dell, G., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language 54,* 228–264.

Scott, R.M., & Wilshire, C.E. (in publication). Lexical competition for production in a case of Broca's aphasia. Converging evidence from three different tasks.

Thompson-Schill, S.L., & Botvinick, M.W. (2006). Resolving conflict: A response to Martin and Cheng (2006). *Psychonomic Bulletin & Review, 13 (3),* 402-408.

Wheeldon, L.R., & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. *Journal of Memory and Language, 33,* 332-356.

Wilshire, C.E., Keall, L.M., Stuart, E.J., & O'Donnell, D.J. (2007). Exploring the dynamics of aphasic word production using the picture–word interference task: A case study. *Neuropsychologia, 45,* 939-953

Wilshire, C.E., & McCarthy, R.A. (2002). Evidence for a context-sensitive word retrieval disorder in a case of nonfluent aphasia. *Cognitive Neuropsychology, 19 (2)*, 165-186.

Wilson, M. (1987). MRC Psycholinguistic Database (Version 2.00) (Software). Informatics Division Science and Engineering Research Council Rutherford Appleton Laboratory Chilton, Didcot, Oxon, OX11 0QX.

Appendix

Table 2

Catch trial target and accompanying pictures for Experiment 1

Target pictures	Accompanying pictures		
Arm	Back	ear	Pipe
bed	Ball	Earth	Plate
Belt	Band	egg	Queen
Blood	Bean	Eye	road
Branch	bear	fish	roof
Bus	Beard	Flag	Root
Chair	Bell	foot	rope
Cheek	Bike	Gate	Rose
Chest(body)	Bird	ghost	School
Chest(treas)	bomb	Girl	shark
Cloud	Bone	Glass	Sheep
Cup	Boot	glove	Shirt
Fence	Bowl	Goat	Shoe
Fly	box	Gold	Skull
Harp	Boy	Grass	Snake
Hat	Brain	hair	Soap
Key	Bridge	Hand	Soup
Leaf	Bull	Heart	Spring
Light	Cake	Heel	Squirrel
Mask	car	Horse	Stage
moose	Cat	Jaw	Straw
cheese	Cave	Knee	Sun
king	Chain	Knife	Swing
tent	Cheque	lamp	sword
Pen	Chin	Leg	Tear
pig	Cliff	Lung	Thumb
Pool	Clock	man	tie
Ring	Clothes	Match	Toe
Skirt	cork	Milk	Tongue
Smoke	Cow	Moon	Train
Snail	Crown	Mouth	Truck
Star	Desk	Nose	Trunk(eleph)
Tooth	Dog	nun	Trunk(tree)
Tray	Door	Nurse	Watch
Tree	dress	Palm	wig
Van	Drum	Pill	Witch

Table 3

Catch trials for the first experimental session for Experiment 1 (in order of presentation)

Trial	First	Second	Target
1	rope	CUP	CUP
2	Girl	Match	pig
3	Bowl	Crown	Tree
4	Milk	Snail	Snail
5	PIPE	DESK	Chest(body)
6	Spring	Chest(treas)	Chest(treas)
7	DOOR	Tear	Tooth
8	Cow	Horse	Mask
9	Boot	king	king
10	Leaf	Eye	Leaf
11	Light	Jaw	Light
12	Flag	Bus	Bus
13	Cliff	Heel	Tray
14	cheese	Goat	cheese
15	Soap	road	Blood
16	Skirt	lamp	Skirt
17	Boy	Root	Arm
18	Band	Sheep	bed
19	Smoke	Rose	Smoke
20	Harp	Truck	Harp
21	Grass	Ring	Ring
22	Cake	Queen	Star
23	Hat	Mouth	Hat
24	Cloud	WATCH	Cloud
25	ear	Tongue	Key
26	Palm	Pool	Pool
27	Beard	PEN	PEN
28	Cheek	ghost	Cheek
29	tent	BIKE	tent
30	dress	Fly	Fly
31	Plate	BELT	BELT
32	moose	sword	moose
33	Fence	Nurse	Fence
34	Skull	Glass	CHAIR
35	Ball	Van	Van
36	Snake	Branch	Branch

Table 4

Catch trials for the second experimental session for Experiment 1 (in order of presentation)

Trial	First	Second	Target
1	Star	Thumb	Star
2	Lung	Trunk(tree)	BELT
3	Arm	foot	Arm
4	bomb	CHAIN	Pool
5	Truck	Harp	Harp
6	pig	Knee	pig
7	BIKE	tent	tent
8	Bean	Bull	Ring
9	WATCH	Cloud	Cloud
10	sword	moose	moose
11	Dog	Hand	king
12	bed	Bone	bed
13	Mouth	Hat	Hat
14	Tray	Sun	Tray
15	Earth	Stage	Branch
16	Blood	Trunk(eleph)	Blood
17	CLOCK	hair	Snail
18	box	Back	PEN
19	ghost	Cheek	Cheek
20	Swing	Witch	CUP
21	Tree	Soup	Tree
22	Mask	Leg	Mask
23	Rose	Smoke	Smoke
24	Nurse	Fence	Fence
25	Eye	Leaf	Leaf
26	Straw	Cave	Fly
27	Cheque	Pill	Chest(treas)
28	Key	Bridge	Key
29	Chest(body)	wig	Chest(body)
30	Bell	Gold	Van
31	CHAIR	Nose	CHAIR
32	Jaw	Light	Light
33	car	glove	Bus
34	lamp	Skirt	Skirt
35	Goat	cheese	cheese
36	Tooth	nun	Tooth

Table 5

Catch trials for the third experimental session for Experiment 1 (in order of presentation)

Trial	First	Second	Target
1	Shoe	Bird	Smoke
2	Pool	Palm	Pool
3	Fly	dress	Fly
4	Bus	Flag	Bus
5	shark	man	tent
6	Moon	Heart	cheese
7	wig	Chest(body)	Chest(body)
8	Van	Ball	Van
9	king	Boot	king
10	Chest(treas)	Spring	Chest(treas)
11	CUP	rope	CUP
12	Ring	Grass	Ring
13	Clothes	cork	Cheek
14	BELT	Plate	BELT
15	bear	fish	Fence
16	Thumb	Star	Star
17	Snail	Milk	Snail
18	Cat	Train	Skirt
19	DRUM	Chin	Leaf
20	Nose	CHAIR	CHAIR
21	Soup	Tree	Tree
22	Bone	bed	bed
23	KNIFE	Squirrel	Light
24	roof	Toe	Hat
25	Bridge	Key	Key
26	PEN	Beard	PEN
27	Sun	Tray	Tray
28	Trunk(eleph)	Blood	Blood
29	Leg	Mask	Mask
30	Knee	pig	pig
31	nun	Tooth	Tooth
32	Branch	Snake	Branch
33	School	tie	Cloud
34	Gate	Brain	Harp
35	foot	Arm	Arm
36	Shirt	egg	moose

Table 6

Catch trial target and accompanying pictures for Experiment 2

Target pictures	Accompanying pictures		
Ant	Back	foot	Plate
Arm	Ball	Fork	Plug
bed	Band	Fridge	Plum
Bee	Bank	Gate	rake
Bird	Barn	Girl	Ring
Bread	beak	Glass	scarf
Bull	bear	globe	School
Bus	Beard	Goal	Sheep
Car	bed	Goat	Shoe
Cat	Beer	Gold	sink
Chest(body)	Bell	Gong	skirt
Corn	Belt	Grapes	Sled
Cow	Bike	Grass	slide
crab	Blood	Harp	Smoke
dove	bomb	hose	Snail
Dress	Bone	House	Snake
Duck	Book	Jaw	Soap
eel	Boot	Key	Sock
Eye	Bow	Kite	Soup
Flute	Bowl	Knife	Spring
Frog	Brick	Lake	squid
glove	Bridge	Lamp	Squirrel
gun	cards	Leaf	Stage
Hand	Cast	Leg	Star
Hat	Cave	Light	Stool
Horse	Chain	Map	Straw
Lion	Chair	Mask	Suit
Milk	Cheese	Match	Sun
Mouse	Cheque	mop	Swan
Pear	Cliff	Neck	Sweat
pie	Clock	Net	Swing
Plane	Clothes	Nose	sword
Seal	Cloud	nun	Tail
Shirt	Comb	box	Tear
Skunk	cork	fish	Thigh
Tree	cot	ghost	Tongue
	Crack	king	Train
	cross	man	Truck
	Crown	nest	Trunk(eleph)
	Cup	road	tusk
	dart	roof	Van
	Desk	rope	Waist
	dice	shark	wall
	Dog	tent	Watch

doll	tie	Wave
Door	wig	Well
Drum	Owl	whale
Egg	Pants	Wine
Fan	Peg	Wing
Fence	Pen	Witch
Film	pig	wolf
Flag	Pipe	worm
Fly	plant	

Table 7

Catch trials for the first experimental session for group 1, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	hose	Plane	Cup	Plane
2	Bell	cards	eel	eel
3	WATCH	Seal	rope	Seal
4	squid	PEN	dove	dove
5	sword	gun	Crack	gun
6	Back	Arm	Lake	Arm
7	Corn	Bone	Tail	Corn
8	Goat	Cow	Bank	Cow
9	bed	DESK	Wave	bed
10	Mouse	Gold	Wine	Mouse
11	School	Net	Tree	Tree
12	Grapes	nest	Pear	Pear
13	Swan	Horse	Fridge	Horse
14	House	Wing	pie	pie
15	Dog	Lion	slide	Lion
16	Witch	Skunk	Cave	Skunk
17	Bread	Cliff	Fork	Bread
18	whale	crab	doll	crab
19	Cat	Fence	sink	Cat
20	Frog	Snake	cot	Frog
21	road	Comb	Bull	Bull
22	Lamp	Chest(body)	nun	Chest(body)
23	Snail	FAN	Bee	Bee
24	Cheque	Shirt	Goal	Shirt
25	Car	Truck	BOOK	Car
26	Key	Hand	Peg	Hand
27	Dress	Suit	box	Dress
28	Harp	Cast	Flute	Flute
29	Duck	ghost	Spring	Duck
30	Beard	Swing	glove	glove
31	Milk	PIPE	beak	Milk
32	Bird	fish	Map	Bird
33	Van	Clothes	Bus	Bus
34	Hat	scarf	dart	Hat
35	Fly	Bow	Ant	Ant
36	Smoke	tusk	Eye	Eye

Table 8

Catch trials for the first experimental session for group 2, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	Train	Plane	Cup	Plane
2	shark	cards	eel	eel
3	bear	Seal	rope	Seal
4	Band	PEN	dove	dove
5	Bridge	gun	Crack	gun
6	Stage	Arm	Lake	Arm
7	Corn	Plum	Tail	Corn
8	Match	Cow	Bank	Cow
9	bed	king	Wave	bed
10	Mouse	Squirrel	Wine	Mouse
11	plant	Net	Tree	Tree
12	Ball	nest	Pear	Pear
13	Sun	Horse	Fridge	Horse
14	Soup	Wing	pie	pie
15	Bowl	Lion	slide	Lion
16	pig	Skunk	Cave	Skunk
17	Bread	Cheese	Fork	Bread
18	Soap	crab	doll	crab
19	Cat	wolf	sink	Cat
20	Frog	tent	cot	Frog
21	Sheep	Comb	Bull	Bull
22	Thigh	Chest(body)	nun	Chest(body)
23	Plate	FAN	Bee	Bee
24	Pants	Shirt	Goal	Shirt
25	Car	Glass	BOOK	Car
26	foot	Hand	Peg	Hand
27	Dress	Flag	box	Dress
28	Straw	Cast	Flute	Flute
29	Duck	Owl	Spring	Duck
30	BELT	Swing	glove	glove
31	Milk	Egg	beak	Milk
32	Bird	Gate	Map	Bird
33	mop	Clothes	Bus	Bus
34	Hat	Cloud	dart	Hat
35	CHAIN	Bow	Ant	Ant
36	Nose	tusk	Eye	Eye

Table 9

Catch trials for the second experimental session for group 1, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	Bus	Train	Peg	Bus
2	Ant	worm	cards	Ant
3	Van	Car	Film	Car
4	Eye	Plug	wig	Eye
5	Flute	DRUM	Light	Flute
6	Crown	Hat	Straw	Hat
7	House	Brick	Shirt	Shirt
8	eel	wall	nun	eel
9	pig	Bird	road	Bird
10	Glass	Duck	Beard	Duck
11	Snake	Fence	Lion	Lion
12	tent	Cat	mop	Cat
13	Sweat	cork	Plane	Plane
14	Witch	Ball	Hand	Hand
15	CHAR	bed	Smoke	bed
16	Trunk(eleph)	Band	Chest(body)	Chest(body)
17	bomb	Kite	gun	gun
18	dove	Owl	Cliff	dove
19	Bull	Match	Net	Bull
20	tie	Dress	CLOK	Dress
21	foot	PIPE	Arm	Arm
22	Cloud	Bread	Gold	Bread
23	Bone	Mouse	School	Mouse
24	Bee	Fly	Tail	Bee
25	Bridge	Cast	Skunk	Skunk
26	Stage	Corn	DOOR	Corn
27	bear	ghost	Horse	Horse
28	Squirrel	Plate	Cow	Cow
29	glove	Well	slide	glove
30	shark	Frog	Bell	Frog
31	Tree	Key	cot	Tree
32	Sled	Milk	CHAIN	Milk
33	Tear	PEN	Seal	Seal
34	pie	Blood	Star	pie
35	Pear	Egg	Girl	Pear
36	fish	rake	crab	crab

Table 10

Catch trials for the second experimental session for group 2, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	Bus	Cheque	Peg	Bus
2	Ant	Ring	cards	Ant
3	Soap	Car	Film	Car
4	Eye	Jaw	wig	Eye
5	Flute	Barn	Light	Flute
6	BIKE	Hat	Straw	Hat
7	tie	Brick	Shirt	Shirt
8	eel	squid	nun	eel
9	Flag	Bird	road	Bird
10	Dog	Duck	Beard	Duck
11	Crack	Fence	Lion	Lion
12	Swan	Cat	mop	Cat
13	Truck	cork	Plane	Plane
14	Leg	Ball	Hand	Hand
15	globe	bed	Smoke	bed
16	Waist	Band	Chest(body)	Chest(body)
17	Shoe	Kite	gun	gun
18	dove	wall	Cliff	dove
19	Bull	wolf	Net	Bull
20	man	Dress	CLOK	Dress
21	king	PIPE	Arm	Arm
22	Soup	Bread	Gold	Bread
23	Sheep	Mouse	School	Mouse
24	Bee	dice	Tail	Bee
25	Goat	Cast	Skunk	Skunk
26	Grapes	Corn	DOOR	Corn
27	Lamp	ghost	Horse	Horse
28	Sun	Plate	Cow	Cow
29	glove	Boot	slide	glove
30	cross	Frog	Bell	Frog
31	Tree	Grass	cot	Tree
32	Beer	Milk	CHAIN	Milk
33	whale	PEN	Seal	Seal
34	pie	Cheese	Star	pie
35	Pear	BOOK	Girl	Pear
36	WATCH	rake	crab	crab

Table 11

Catch trials for the third experimental session for group 1, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	Bow	Fridge	Corn	Corn
2	Seal	Tear	Bank	Seal
3	Plug	Eye	rope	Eye
4	Bone	Plate	Mouse	Mouse
5	Chest(body)	Trunk(eleph)	Spring	Chest(body)
6	Sled	Cast	Milk	Milk
7	gun	KNIFE	Stage	gun
8	Key	Tree	box	Tree
9	Shirt	hose	Cliff	Shirt
10	Glass	Cheque	Duck	Duck
11	Cow	fish	Straw	Cow
12	Skunk	Beard	nest	Skunk
13	Cloud	Film	Bread	Bread
14	tent	sink	Cat	Cat
15	Stool	Crack	bed	bed
16	Squirrel	Bell	Frog	Frog
17	Train	Wave	Car	Car
18	Dog	rake	Bird	Bird
19	worm	Bee	doll	Bee
20	Hand	Witch	roof	Hand
21	Soup	Pear	Clothes	Pear
22	Mask	School	Hat	Hat
23	skirt	Map	Dress	Dress
24	Truck	Bus	Wing	Bus
25	Plane	Sweat	Goal	Plane
26	Horse	Goat	Smoke	Horse
27	Bridge	glove	tusk	glove
28	Arm	Nose	Lake	Arm
29	Match	Bull	dart	Bull
30	Blood	eel	Star	eel
31	Swan	dove	PIPE	dove
32	Snail	Ant	Well	Ant
33	crab	squid	Fork	crab
34	Gong	Flute	Swing	Flute
35	wall	pie	beak	pie
36	Lion	Owl	FAN	Lion

Table 12

Catch trials for the third experimental session for group 2, Experiment 2 (in order of presentation)

Trial	Bounce	Flash	Stationary	Target picture
1	Soup	Fridge	Corn	Corn
2	Seal	wolf	Bank	Seal
3	Tongue	Eye	rope	Eye
4	Snake	Plate	Mouse	Mouse
5	Chest(body)	Neck	Spring	Chest(body)
6	Wine	Cast	Milk	Milk
7	gun	Cup	Stage	gun
8	Leaf	Tree	box	Tree
9	Shirt	Boot	Cliff	Shirt
10	Sheep	Cheque	Duck	Duck
11	Cow	man	Straw	Cow
12	Skunk	Dog	nest	Skunk
13	Egg	Film	Bread	Bread
14	bear	sink	Cat	Cat
15	globe	Crack	bed	bed
16	cross	Bell	Frog	Frog
17	Soap	Wave	Car	Car
18	Flag	rake	Bird	Bird
19	hose	Bee	doll	Bee
20	Hand	Back	roof	Hand
21	BOOK	Pear	Clothes	Pear
22	BIKE	School	Hat	Hat
23	Sun	Map	Dress	Dress
24	dice	Bus	Wing	Bus
25	Plane	Van	Goal	Plane
26	Horse	Lamp	Smoke	Horse
27	Sock	glove	tusk	glove
28	Arm	king	Lake	Arm
29	pig	Bull	dart	Bull
30	whale	eel	Star	eel
31	Brick	dove	PIPE	dove
32	Ring	Ant	Well	Ant
33	crab	WATCH	Fork	crab
34	Barn	Flute	Swing	Flute
35	Grapes	pie	beak	pie
36	Lion	Bowl	FAN	Lion